Working Meeting to Address HIPPI-6400 Issues Only
Tuesday, Wednesday and Thursday, January 7-9, 1997
Phoenix, Arizona

1. Opening remarks and introductions

The Chairman, Don Tolmie of Los Alamos National Laboratory, opened this HIPPI-6400 meeting and thanked Chris Olson and Lockheed Martin for hosting this meeting. This group is constituted as both the HIPPI special working group (SWG) under X3T11, and the HIPPI Networking Forum (HNF) - Technical Committee (TC).

Don then lead a round of introductions. The list of attendees is at the end of these minutes.

2. Review / modify the draft agenda

The draft agendas were available on the web prior to the meeting. Hard copies were available at the meeting. James Hoffman of Los Alamos volunteered to take the meeting minutes. The AMP connector proposal, 5.1.3, was removed as AMP did not finish their proposal in time for the meeting. Berg presented its new Jackscrew connector (5.1.3). Dean Liberty's PH comments were reviewed (5.7). A discussion on further splitting the document was also added (5.8). The group discussed broadcast techniques for 6400 (5.9). The Port_Teardown state table was reviewed (5.10).

3. Review minutes of previous meetings

3.1 December 2-3, 1996, Minneapolis, MN

The Minneapolis minutes were reviewed. Roger Ronald moved and Jeff Young seconded to approve the minutes as written. Approved unanimously.

3.2 Review action items from Minneapolis meeting

1. Greg Chesson and Fred Templin to provide ARP text for inclusion in HIPPI-6400-SC, and specify effects on bridging. (Carryover, see 5.9)
2. Greg Chesson to review counter size of SuMAC Retransmission_Error counter and the need for both contiguous retransmission and total retransmission error counters. (Done, see 4.1)
3. Hansel Collins to investigate ways of specifying interoperability voltages by detailing components, parasitics, and parameters at the driving and receiving ends. (Done)
4. Greg Chesson to register the EtherTypes with INA / Xerox. (In process)
5. Greg Chesson to have Art Beckman look into getting a 12-bit group of ULA's for HNF. (In process)
6. Greg Chesson to present text to reflector to describe RTR setup using the Persistent bit. (Carryover)
7. Don Tolmie to review crossing-in-the-night Port_Teardown and End operation cases and review the solution with Wally St. John. (Done)
8. Hansel Collins to find safe operation range for Input High Voltage and Input Low Voltage. (Done)
9. Henry Brandt and Hansel Collins to collect values for completion of copper interface specifications. (In process)
10. John Ellis to determine retention force of connector. (Overcome by events)
11. John Ellis to redo values for near end crosstalk. (Done)
12. Michael Karg to find the international cable specification for plenum rating. (Done)
13. Greg Chesson and Don Sanders to have SGI review the connector layout as it relates to the SuMAC chip. (Carryover)
14. Greg Chesson to see if we can make 8.7 and 8.1 in -ST consistent, i.e., 0 or -1 for S_id and T_id don't care values. (Done)
15. Greg Chesson and Don Sanders to check all error states, e.g., what happens if a command is corrupted or lost. (In process)
16. Roger Ronald to draft -ST over HIPPI-FP so that we can see if it obviates the need for HIPPI-MP. (Done)
17. Roger Ronald to update HIPPI-6400-SC Rev 0.6 with changes agreed to at the Minneapolis meeting. (Done)
18. Don Tolmie to update HIPPI-6400-ST Rev 0.1 with the changes agreed to at the Minneapolis meeting. (Done)

19. Don Tolmie to update HIPPI-6400-PH Rev 0.85 with the changes agreed to at the Minneapolis meeting. (Done)

4. Review document changes since last meeting

These sections present new or controversial changes to each of the documents. The changes to the document prior to the meeting can be found in the corresponding change list at the front of each document.

4.1 HIPPI-6400-PH, Rev 0.9

Fixed the Introduction and Scope first bullet on retransmission. 6400 retransmits to correct transmission errors as they occur.

Craig Davidson of E-Systems received an action item to add text to 4.10 describing the 16-bit optical interface.

The note at the end of 6.2 describing VC sizing was revised. The description of VC3's size restriction was clarified.

In 7.1 added that the ULA's determine Originating Source and Final Destination. The group asked for more description for Figure 11. Fred Templin volunteered to provide text describing the I/G and U/L bits in Figure 11.

Fulfilling his action item for 8.4, Greg Chesson noted that SGI has a counter called "Retransmissions" which sums the RSEQ_Missing_Error and RSEQ_Out_Of_Range_Error counters. He did not feel that any specific value of Retransmissions should cause shutdown, (i.e., only failed successive retransmissions should cause shutdown.)

In 11.2, the timeout mentioned in 11.2 was named "training timeout". The sections title was changed to "Training sequence errors".

Figure 16 – Initialize and Link Reset operations was thoroughly reviewed and modified. The "Start Hold-off timer" sequences were moved to the bottom box. The bottom box now points to the top dotted box. Power-On now causes an Initialize. The training timeout starts in the same box as "Reset local state" for both Reset and Initialize. The CLOCK_2 reference becomes an "activity monitor" so it can apply to both optical and copper.

In 12.3, the paragraph about "External administrative action..." needs to be reworded to fit with the new text above.

The text describing CLOCK_2 in clause 14 was reviewed and revised. Once again it was noted that CLOCK_2 serves as both an activity monitor for the link and as a free-running clock for 16-bit implementations. The group suggested the use of the word "activity monitor" to tie in with Figure 16. The activity monitor will be different for optical systems and may be incorporated into the open fibre control system.

4.2 HIPPI-6400-SC, Rev 0.7

The group reviewed HIPPI-6400-SC changes cover to cover with minor revisions.

The group examined Table 7 and suggested some formatting and consistency changes.

The Reset_Response command was removed as it received only opposition on the reflector.

Fred Templin took an action item to make sure that RFC 1700 lists all reserved ULA addresses.

4.3 HIPPI-ST, Rev 0.2

The group spent a day reviewing the ST document and changes abound throughout the document. As always, check the change bars of ST 0.3 for exact changes.

HIPPI-ST needs a reference to the EtherType numbers like HIPPI-6400-PH.

Figure 2 doesn't relay much information and the group asked for simplification.

The note of performance increase when data and control channels are separated was removed as it implies an unknown.

Figure 4 was heavily reviewed and revised. Various terms and the lower part of the drawing added too much confusion for many readers. The text below the figure was also reworded.

The group declared a global change of minimum STU, buffer, and block size to be greater than or equal to 256 bytes.

The STU-Size, 4.3.5, was completely redone and renamed to Max-STU to remove confusion with STU sizing elsewhere in the document.

Greg Chesson received an action item to reword section 4.3.6 on Slots to incorporate the idea of a separate data handling unit that does not require any slots unless interrupt or notify is set on the corresponding DATA operation.

A new section was added to 4.4 called Block size. The block size is now as an integral power of two larger than 256 bytes. The field is decoded where the parameter indicates the power of two of the block size.

In 4.4.3, block ordering in the second paragraph was changed. An Out-of-Order bit will be set in the RQP to denote if an endpoint can send or receive blocks out of order. Out of order transmission and reception implies retransmission capability.

Many members of the group had a tough time with tiling considerations. A new section may be added to the next document displaying and detailing the use of Max-STU, block sizes, buffers sizes, (e.g., source STU sizing restrictions).

In 6.1, the group requested a new format where each capability of a parameter is separately bulleted.

The D_Port was changed to R_Port (same change for D_id) to indicate receiver port. The words sender and receiver indicate the sender and receiver of an operation and don't relate to the Final Destination or Originating Source.

Each of the individual operations were reviewed with editing revisions to comply with other recent changes.

The error processing clause was moved around with minor editing. The clause was made more consistent with regards to error logging for each error.

5. New Items

5.1 HIPPI-6400-PH copper

Chris Olson noted that thin film chip packaging would allow placement of the HIPPI-6400 equalization network in a 1 inch by 1 inch chip.

Don started the copper section by reading an e-mail from Chuck Brill of AMP. AMP planned to bring a connector proposal to the January meeting, but time limitations hindered their proposal. They will support whatever choice the committee makes.

The group reviewed clause 15 and fixed or removed outdated parameters and text. In particular the group wants better separation between cable and connector specifications. They decided that the backshell will not be connected to any cable shield or pins. The group also decided that the Reserved pins will now be No Connect.

Although informally selected previously, Roger Ronald moved and Henry Brandt seconded to select the Berg Micropax 100-position connector for the HIPPI-6400 16-bit copper interface standard. Motion passed, 9 for, 0 opposed, and 1 abstention.

John Ellis received an action item to find the correct part number for this connector. Don Tolmie elaborated on the Berg's licensing stance for this connector. Another vendor may produce the connector with access fee to all of Berg's connector drawings for a $35,000 license fee plus royalties.

The group had a question on electrostatic buildup in these new cables, especially with the lower dielectric. The group may use grounding covers on the end of each cable to alleviate this problem.

Finally, the bulkhead to board connection was reviewed. Previously, the PH specification called for a straddle mount connector, but all companies requested a right angle connector instead. The Berg connector part number will specify a right angle connector, but individual companies may use a straddle mount connector if they choose.

5.1.1 Hansel Collins' proposal

Hansel Collins walked through his electrical specification which was available on the web before the meeting. Each parameter was discussed and

some revisions were made. Hansel was added to Henry Brandt's action item to collect the copper specifications for completion. Part of the action item is also to decide which parts of Hansel's presentation should be incorporated into the PH specification. The purpose of Hansel's specification is to provide other vendor's with all the external SuMAC properties up to the bulkhead.

Hansel asked James Hoffman to place copies of the Widmer worst case frequency pattern on the web page so vendors can test their equipment.

The group questioned Hansel about the SuMAC crosstalk precautions. Hansel said they are using strict spacing design guidelines that remove crosstalk from the outset.

Hansel was asked if the destination 16-bit end could add a blocking capacitor for better ESD characteristics. Unfortunately, this adds to the line charge time and hurts performance.

5.1.2 Grounding the overall shield

Don Tolmie started the shield conversation by reading Gary Vincent's email January 8th reflector comments. Vincent's comments call for grounding at both ends which would work for setups going into a single installation, but HIPPI-6400's 50 meter capability allows it bridge between different installations which may have very different potential levels.

For an AC coupled ground the group proposed to use an asymmetrical cable ground with grounds on pins 51 and 100 at one end and capacitively coupled at the other end. Greg Chesson moved, and John Ellis seconded, to accept the asymmetrical cable. The motion passed, 9 for, 0 opposed, and 3 abstentions.

5.1.3 Jackscrew connector

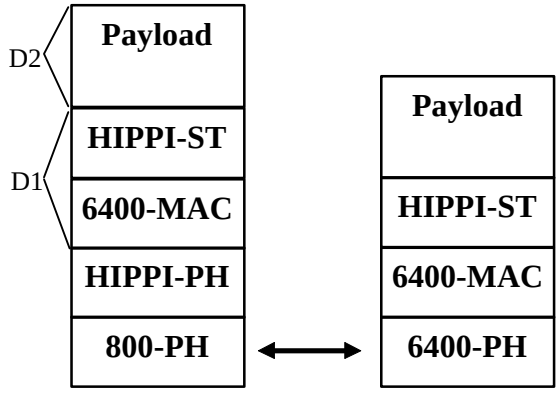
John Ellis of Berg brought in a jackscrew version of the Micropax 100-pin connector. The jackscrew configuration allows a connector to have a 45 degree exit from the bulkhead. The 45 degree exit angle allows companies to meet requirements for separation between the bulkhead and the cabinet back (E-Systems quoted a requirement of eight inches.) Craig Davidson moved, and Hansel Collins seconded, to use a connector with jackscrews. Motion passed, 9 for, 0 opposed, and 3 abstentions.

5.2 HIPPI-ST Annex A – mapping to HIPPI-6400-PH

The group reviewed Don's first shot at showing the mapping between HIPPI-ST above HIPPI-6400-PH. The group did a read through with a few revisions.

5.3 HIPPI-ST Annex B – mapping 6400 to 800

Craig Davidson's proposal was available on the web before the meeting, and hard copies were available at the meeting. The figure below gives a brief overview of the proposed protocol stack when carried between HIPPI-PH (800) and HIPPI-6400-PH.



The first question is whether the D1 space should cover the two headers or whether everything should be part of D2. Greg Chesson felt that SGI's current OS-bypass protocol does use the D1 field. Jeff Young of Cray said that Cray may need the D1 area. If the D1 is kept it can be placed in the first burst (short burst) of a HIPPI-800 message. This will require the last burst to be full size and may require padding. If everything is considered part of D2 the last burst may be short. Jeff Young took an action item to review Cray driver requirements involving D1, D2 and burst boundaries. Superseding that action item, Roger Ronald plans to email the reflector with the same question to all driver implementors.

First off, the group noted that some techniques used for the first translator may be specific to E-Systems needs, but should pave the way for a stronger standard.

The group reviewed all the figures and the two examples with minor corrections. They also asked for a more compact figure representation.

The translator has the capability of fragmenting and reassembling transfers from 800 to 6400 to work with the small buffers of the translator and the low interrupt cycle present on many 800 devices. Hence, lots of small 6400 messages may be assembled into large 800 packets in the translator. The translator can fragment a single 800 CTS into multiple 6400 CTS'es to decrease 6400 congestion. The translator also operates in a bypass mode where incoming CTS'es are untouched.

One concern that was touched was error recovery through the translator. Obviously, the translator cannot retransmit old Blocks, but it can pass the flawed data to the end point which can detect the error and retransmit on a Block basis.

5.4 HIPPI-ST Annex C – Examples

Annex C has hung behind the ST development for a while since ST continues to change at a rapid pace. The group decided that Annex C should contain multiple examples showcasing various ST attributes as they are nailed down. Possible examples include: basic transfer, RTR example, direct addressing example with persistence on, and a translator example.

5.5 HIPPI-6400-PH service interface

The group reviewed Don's first cut at expanding the service interface. A few comments were made on the exact use and breakdown of signals. The text shall be inserted into the next PH draft.

5.6 HIPPI-6400-PH training sequence disparity

The training sequence disparity as shown in Figure 14 and 15 was discussed. The group agreed that the training detect portion of the sequence shall bring the running disparity to 0 or +1. The tail of training sequence causes a -1 disparity resulting in a final disparity of 0 or -1 which is then cleared to 0. Previously the sequence resulted in -1 or -2 and was cleared to 0.

5.7 Dean Liberty's PH comments

Dean Liberty of IBM sent Don some comments on the introduction sections of HIPPI-6400-PH. Running out of time the group left the inclusion of the items to the editor with review at the next meeting.

5.8 Splitting PMD out of the document

Don Tolmie led a discussion about splitting the PMD (16 bit copper and 8 bit optical interfaces) out of the HIPPI-6400-PH document. Though it might allow easier reading of the split documents, the group did not think the time exists in the tight schedule to split the documents. This topic may be reviewed at future meetings.

5.9 6400 broadcast discussion

One important aspect of 6400 is using standard techniques for IP communication. To provide standard ARP and ICMP services the group would like to implement a broadcast mechanism in HIPPI-6400. No formal presentations were made but the ideas are presented below.

The group would like to use a broadcast server that receives and propagates all messages designated by ULA "FFFFFF". Using a single broadcast server on the network eliminates fragmented broadcast lists between multiple servers. The group noted that hosts should bear as much of the setup work as possible (keeping switches simple in this regard).

The following tasks are required for a single broadcast server protocol:

- switches need to allow hosts to register for multiple addresses (normal and broadcast)
- host broadcast servers need to dynamically detect host addresses
- detection of down broadcast servers and activation of one new broadcast server.

Other items that should be considered include:

- The broadcast layer should remain separate from the IP layer so that any protocol can use broadcast.
- Upward mobility should be maintained for Ethernet translation (may require large tables).
- Consistency with current documentation and implementations on HIPPI-800 networks (if possible).

Fred Templin presented a first stab at resolving some of the broadcast issues. First he noted that we are doing all-multicast and not broadcast as hosts not running broadcast routines cannot broadcast or

receive broadcast. A summary of actions taken for broadcast appear below:

1. Host begins neighbor discovery using EXCHANGE_ELEMENT_FUNCTION admin command. Two new element types include: Non-switch broadcast server capable Element and Broadcast Switch Element. Once a switch is found, ULA's are requested.
2. Host issues ULA_REQUEST for ULA's.
3. Switches give ULA's (ULA_RESPONSE) and if no broadcast server is active, give out broadcast ULA to broadcast server capable element.
4. All hosts (except broadcast server) begin issuing periodic "ICMP ping" to broadcast server to update broadcast tables.
5. Broadcast server listens on broadcast ULA for "ICMP pings" and responds.
6. Actual broadcast requests are sequentially unicast to all hosts in broadcast table.

Issues that need to be refined include:

- At startup or when a broadcast server crashes, who becomes the new broadcast server.
- Possible periodic ping to broadcast server to detect a crash.
- Interrelation with broadcast capable switches.
- Restrict broadcast to only certain VC's.

Various people agreed to discuss the initial ideas over the reflector for further development.

5.10 Port_Teardown state table

Don presented the Port_Teardown state table that has been in development between Don Tolmie and Greg Chesson. After learning the new state format, the group applauded the design benefits of the format. The group will finish out the state table before inserting it into an normative annex of ST. Greg Chesson noted that SGI is developing similar state tables for all the ST operations.

6. Future meeting schedule

6.1 February 3-4, 1997, San Jose, CA

During the X3T11 December plenary week , the following HIPPI meetings are scheduled:

Monday, February 3 -
9 AM - 9 PM — HIPPI-6400

Tuesday, February 4 -
8 AM - 9 AM — HNF Plenary
9 AM - 2 PM — HIPPI-TC General and -6400
3 PM - 5 PM — HIPPI-6400 Optical (with connector review and selection)
6 PM - 9 PM — HIPPI-6400 General

The location is the Red Lion Hotel, 2050 Gateway Place, San Jose, CA 95110. Bob Snively and Sun Microsystems are the host (See the meeting announcement on the web page at http://www.cic-5.lanl.gov/~det/ for further details.)

6.2 March 4-6, 1997, San Jose, CA

During the March interim week , the following HIPPI meetings are scheduled. **Note that an extra day has been added.**

Tuesday, March 4 -
1 PM - 9 PM — HIPPI-6400 Copper

Wednesday, March 5 -
8 AM - 9 PM — HIPPI-6400
Thursday, March 6 -
8 AM - 5 PM — HIPPI-6400

The location is San Jose, CA. Berg is the host. At this time the meeting hotel was unknown, but will be announced as soon as it is picked. The meeting announcement will be on the web page at http://www.cic-5.lanl.gov/~det/.

6.3 Future meeting dates and locations

The following 1997 X3T11 plenary week dates are firm. Recent changes to this list are underlined to make them easier to find.

1997 –

Feb 3-4	Plenary	San Jose, CA	Sun
Mar 4-6	Interim	San Jose, CA	Berg
Apr 7-8	Plenary	Palm Springs, CA	Brocade
May 13-15	Interim	Mt. View, CA	SGI
Jun 9-10	Plenary	Seattle, WA	Boeing
July 8-10	Interim	Minneapolis, MN	Cray
Aug 4-5	Plenary	Honolulu, HI	Hitachi
Oct 6-7	Plenary	Tucson, AZ	FSI
Dec 8-9	Plenary	Orlando, FL	DPT

Los Alamos and E-Systems have agreed to host interim meetings (September in Albuquerque and November in Dallas) if needed. A decision will be made at a later meeting.

The 1998 schedule is less firm, but here is what is currently being considered by X3T11 for the plenary meetings. Question marks note the ones that are still in question. Hopefully HIPPI-6400 will be far enough along that we will not need interim working meetings.

1998 –

Feb 9-10	Plenary	San Diego	Qlogic
Apr 20-21	Plenary	Palm Springs, CA	Brocade
Jun 8-9	Plenary	St. Petersburg Beach, FL	AMP
Aug 10-11	Plenary	??	??
Oct 5-6	Plenary	Tucson, AZ (?)	FSI (?)
Dec 7-8	Plenary	Ft. Lauderdale, FL	Adaptec

7. Review action items

All of the following action items apply to HIPPI-6400.

1. Greg Chesson and Fred Templin to provide ARP text for inclusion in HIPPI-6400-SC, and specify effects on bridging.
2. Greg Chesson to register the EtherTypes with INA/Xerox.
3. Greg Chesson to have Art Beckman look into getting a 12-bit group of ULA's for HNF.
4. Greg Chesson to present text to reflector to describe RTR setup using the Persistent bit.
5. Henry Brandt and Hansel Collins to collect values for completion of copper interface specifications.
6. Greg Chesson and Don Sanders to have SGI review the connector layout as it relates to the SuMAC chip.
7. Greg Chesson and Don Sanders to check all error states, e.g., what happens if a command is corrupted or lost.
8. Craig Davidson to add text to 4.10 of PH describing the 16-bit optical interface.
9. Craig Davidson to provide text describing how HIPPI-ST Sync values are chosen.
10. Greg Chesson to provide new text for the HIPPI-ST Notify and Interrupt flag bits.
11. Greg Chesson to provide new text for HIPPI-ST Op Retry and error processing clause.
12. Fred Templin to provide text describing the I/G and U/L bits in Figure 11 of PH.
13. Fred Templin to make sure that RFC 1700 lists all of the reserved ULA addresses used in HIPPI-6400.
14. Roger Ronald to email the reflector with burst boundary question for FP encapsulation of ST.
15. John Ellis to send the part number for the 100-pin Berg connector to Don Tolmie for inclusion in the document.
16. James Hoffman to make Al Widmer's worst-case bit patterns available on the HIPPI web page.
17. Jeff Young to review Cray driver requirements involving HIPPI-800 D1, D2 and burst boundaries.
18. Greg Chesson to reword section 4.3.6 on Slots and to create a section on Notify and Interrupt.
19. Roger Ronald to update HIPPI-6400-SC Rev 0.7 with changes agreed to at the Phoenix meeting.
20. Don Tolmie to update HIPPI-6400-ST Rev 0.2 with the changes agreed to at the Phoenix meeting.
21. Don Tolmie to update HIPPI-6400-PH Rev 0.9 with the changes agreed to at the Phoenix meeting.

8. Adjournment

The group adjourned at 5 pm. The thought of the meeting for this month is: "The more we disagree, the greater the chance that one of us is right."

9. Attendance

John Ellis	Berg Electronics	717-938-7512	Ellis825@aol.com
Robert Hooley	Berg Electronics	717-938-7851	robhooley@aol.com
Ralph Papa	Berg Electronics	717-459-5400	
Rich Turner	C&M Corporation	860-779-4225	rmt@cm-corp.com
Michael Karg	Cable Design Technologies	508-791-3161	mikekarg@montrose-cdt.com
Jeff Young	Cray Research	612-683-5536	jsy@cray.com
Michael McGowen	Essential Communications	505-344-0080	mikemc@esscom.com
Randy Hardy	Harris	407-984-5654	rhardy@harris.com
Francois Gaullier	Hewlett Packard	+33 4-7614-5181	Francois_Gaullier@hp-france-om1.om.hp.com
Henry Brandt	IBM	914-433-7684	hrap@vnet.ibm.com
Chris Olson	Lockheed Martin	602-925-7275	chris.e.olson@lmco.com
Gene Dornhoff	Los Alamos National Lab	505-667-8255	agd@lanl.gov
James Hoffman	Los Alamos National Lab	505-665-8557	jamesh@lanl.gov
Don Tolmie	Los Alamos National Lab	505-667-5502	det@lanl.gov
Fred Templin	NASA Ames	415-604-6514	templin@nas.nasa.gov
Craig Davidson	Raytheon E-Systems	972-205-4166	davidson@esy.com
Roger Ronald	Raytheon E-Systems	972-205-8043	rronald@esy.com
Greg Chesson	Silicon Graphics	415-933-3496	greg@sgi.com
Hansel Collins	Silicon Graphics	415-933-2921	hac@sgi.com
Bob Newhall	Silicon Graphics	415-933-5424	newhall@asd.sgi.com
Al Kelley	Tensolite	800-458-9960	kelley@tensolite.com
Bill Kuypers	Tensolite	800-458-9960	bkuypers@tensolite.com
Christine Foster	W. L. Gore & Associates	302-368-2575	cfoster@wlgore.com
Brad Knott	W.L. Gore & Associates	602-438-6916	bknott@wlgore.com